

Filtering in Nonlinear Time Delay Systems

THOMAS K. YU, JOHN H. SEINFELD, AND W. HARMON RAY

Abstract—Linear and nonlinear (extended Kalman–Bucy) filters are derived for systems governed by coupled partial and integro-differential equations. The framework used is sufficiently general that filters for 1) lumped parameter systems having multiple time varying or constant time delays, 2) coupled lumped and hyperbolic distributed parameter systems, and 3) lumped parameter systems with functional time delays, evolve as special cases. Although the filtering equations are the final result, the corresponding smoothing equations are developed as well. The performance of the filter is illustrated through application to a well stirred chemical reactor with external heat exchange.

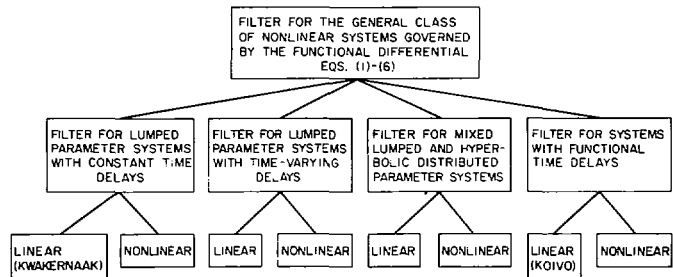


Fig. 1. Classes of filters derived in this paper.

I. INTRODUCTION

RECENTLY there has been interest in filtering for systems described by partial differential equations, [1]–[17], and for systems described by functional differential equations such as those containing time delays (e.g., [18]–[22]). At this time filters for linear and nonlinear distributed parameter systems, linear systems with constant time delays, and linear systems with functional time delays are available. However, filters have not heretofore been available for nonlinear lumped parameter systems containing time delays, either linear or nonlinear lumped parameter systems having time varying delays, or mixed lumped and distributed systems. In this paper, within a single framework we obtain new filters for the following classes of systems:

- 1) nonlinear lumped parameter systems containing multiple constant and time-varying delays;
- 2) mixed nonlinear lumped and hyperbolic distributed parameter systems; and
- 3) nonlinear systems with functional time delays.

Several known [18],[20] and new *linear* filters evolve as special cases of the more general nonlinear results. Fig. 1 illustrates the classes of systems for which filters are derived in this paper.

As in the case of filtering in finite dimensional (lumped parameter) systems, the key mathematical problem in deriving filters for infinite dimensional (distributed parameter) systems is proper definition of the noise processes. Comparatively little work has been published

on mathematically rigorous approaches to the distributed parameter filtering problem, e.g., see Falb [2], Curtain and Falb [23], Kushner [8], and Bensoussan [16], the latter work being a rather comprehensive treatment of filtering in linear distributed parameter systems. Most of the reported work on distributed parameter filtering (as in the case of lumped parameter filtering) has been based on the formal approach of representing the dynamical system as a deterministic partial differential equation forced by a stochastic process with zero mean. Linear formulations of this type have been used by Tzafestas and Nightingale, [5]–[7], [13], Thau [3], Meditch [9], and Sakawa [14]. Nonlinear filtering results based on purely formal approaches have been presented by Tzafestas and Nightingale [7], Lamont and Kumar [17], Seinfeld *et al.* [11], and Hwang *et al.* [12].

Our objective in this paper is to derive approximate nonlinear filters for a wide class of time delay and functional differential systems. Because of the nonlinear character of the problem and the unavailability of rigorous mathematical results applicable to this problem, we have by necessity, adopted a purely formal approach. Essentially, we recast the filtering problem as a deterministic optimal control problem, which we solve by classical techniques to obtain a two point boundary value problem (TPBVP). Decoupling of the TPBVP leads to the desired filtering equations. Although the requisite theory of stochastic differential equations in Hilbert space is not available for the class of problems we consider here, the validity of our results is, in some sense, confirmed by the fact that the two previously derived linear filters of Kwakernaak [18] and Koivo [20] are special cases of the more general class of filters derived here (see Fig. 1). In summary, then, we wish to obtain approximate extended Kalman–Bucy filters for the class of systems shown in Fig. 1.

We begin by formulating the problem in a single framework general enough to include the types of systems depicted in Fig. 1 as special cases. We then present the

Manuscript received January 18, 1973; revised July 26, 1973. Paper recommended by H. W. Sorenson, Chairman of the IEEE S-CS Estimation and Identification Committee.

T. K. Yu is with the Aerospace Group, Missile Systems Division, Hughes Aircraft Company, Canoga Park, Calif.

J. H. Seinfeld is with the Department of Chemical Engineering, California Institute of Technology, Pasadena, Calif. 91109.

W. H. Ray is with the Department of Chemical Engineering, State University of New York at Buffalo, Buffalo, N.Y. 14214.

derivation of the filter. Finally, we illustrate the computational application of the filter through estimation of the temperatures in a chemical reactor-heat exchanger system.

II. FORMULATION OF THE PROBLEM

Let us consider the problem of filtering for the class of systems governed by the coupled ordinary and partial functional differential equations

$$\dot{x}(t) = f(x(t), z(r_1, t), \dots, z(r_\beta, t), t) + \int_0^1 K(z(r, t), r, t) dr + \xi(t) \quad (1)$$

$$z_t(r, t) = -M(r, t)z_r(r, t) + g(z(r, t), r, t) + \zeta(r, t) \quad (2)$$

defined for $t \geq 0$ on the normalized spatial domain $r \in [0, 1]$. $x(t)$ and $z(r, t)$ are n_1 - and n_2 -dimensional state vectors, respectively, and $\xi(t)$ and $\zeta(r, t)$ are zero-mean random processes with arbitrary statistical properties. z_t and z_r denote $\partial z / \partial t$ and $\partial z / \partial r$, respectively. Observations of the system consist of the n_3 -dimensional vector $y(t)$, related to the states by

$$y(t) = h(x(t), z(r_1^*, t), \dots, z(r_\gamma^*, t), t) + \int_0^1 H(z(r, t), r, t) dr + \eta(t) \quad (3)$$

where $\eta(t)$ is a zero-mean measurement error with arbitrary statistical properties and $0 < r_1 < \dots < r_\beta \leq 1$ and $0 < r_1^* < \dots < r_\gamma^* \leq 1$. Thus, the observations can consist in general of the lumped parameter state $x(t)$ and the distributed parameter state $z(r, t)$ at γ measurement locations and integrated over the spatial domain of the system. Initial conditions for (1) and (2) are

$$x(0) = x_0 \quad (4)$$

$$z(r, 0) = z_0(r). \quad (5)$$

The boundary condition at $r = 0$ for (2) is

$$z(0, t) = b(x(t)). \quad (6)$$

We shall now show that by appropriate modification of the system (1)–(6), four important classes of time delay and mixed lumped and distributed parameter systems result.

Nonlinear Systems with Multiple Constant Time Delays

The system (1)–(6) can be reduced to the following nonlinear lumped parameter system containing multiple constant time delays:

$$\dot{x} = f(x(t), x(t - \alpha_1), \dots, x(t - \alpha_\beta), t) + \xi(t) \quad (7)$$

$$y(t) = h(x(t), x(t - \alpha_1^*), \dots, x(t - \alpha_\gamma^*), t) + \eta(t) \quad (8)$$

$$x(t) = \phi(t), \quad -\alpha_{\max} \leq t \leq 0$$

$$\alpha_{\max} = \max(\alpha_\beta, \alpha_\gamma^*) \quad (9)$$

where $0 < \alpha_1 < \dots < \alpha_\beta$ and $0 < \alpha_1^* < \dots < \alpha_\gamma^*$. This can be done by setting $K = H = g = \zeta = 0$, $b(x(t)) = x(t)$, (and hence, $n_2 = n_1$), $M(r, t) = \alpha_{\max}^{-1}$, $r_i = \alpha_i / \alpha_{\max}$, $r_j^* = \alpha_j^* / \alpha_{\max}$, and $z(r, 0) = \phi(-r\alpha_{\max})$. Then, $z(r_i, t) = x(t - \alpha_i)$ and $z(r_j^*, t) = x(t - \alpha_j^*)$. In the formulation (7)–(9) there are β constant time delays in the state equation and γ constant time delays in the observation equation. These delays need not be equal.

Nonlinear Systems with Multiple Time-Varying Delays

The system (1)–(6) can be reduced to the following nonlinear lumped parameter system containing multiple time-varying delays:

$$\dot{x} = f(x(t), x(t - \alpha_1(t)), \dots, x(t - \alpha_\rho(t)), t) + \xi(t) \quad (10)$$

$$y(t) = h(x(t), x(t - \alpha_1^*(t)), \dots, x(t - \alpha_\omega^*(t)), t) + \eta(t) \quad (11)$$

$$\dot{\alpha}_i(t) < 1, \quad i = 1, 2, \dots, \rho \quad (12)$$

$$\dot{\alpha}_j^*(t) < 1, \quad j = 1, 2, \dots, \omega \quad (13)$$

$$x(t) = \phi(t), \quad -\alpha_{\max} \leq t \leq 0 \quad (14)$$

$$\alpha_{\max} = \max(\alpha_1(0), \dots, \alpha_\rho(0), \alpha_1^*(0), \dots, \alpha_\omega^*(0)). \quad (15)$$

To do so we set $K = H = g = \zeta = 0$, $\beta = \gamma = 1$, $r_1 = r_1^* = 1$, $b(x(t)) = [x^T(t), x^T(t), \dots, x^T(t)]^T$, an $n_2 = (\rho + \omega)n_1$ - dimensional vector consisting of $\rho + \omega$ identical vector elements $x(t)$, $M(r, t) = [M_{ij}(r, t)]$ an $n_2 \times n_2$ matrix with $n_1 \times n_1$ matrix components M_{ij} defined by

$$M_{ij} = \begin{cases} 0, & i \neq j \\ \frac{1 - r\dot{\alpha}_i}{\alpha_i} I, & i = 1, 2, \dots, \rho \\ \frac{1 - r\dot{\alpha}_i^*}{\alpha_i^*} I, & i = \rho + 1, \dots, \rho + \omega. \end{cases} \quad (16)$$

Also we let the n_2 -dimensional vector $z(r, t) = [z_1^T(r, t), \dots, z_\rho^T(r, t), z_1^{*T}(r, t), \dots, z_\omega^{*T}(r, t)]^T$ where each $z_i(r, t)$ or $z_j^*(r, t)$ is an n_1 -dimensional vector, and set $z_i(r, 0) = \phi(-r\alpha_i(0))$, $z_j^*(r, 0) = \phi(-r\alpha_j^*(0))$. Then $z_i(1, t) = x(t - \alpha_i(t))$ and $z_j^*(1, t) = x(t - \alpha_j^*(t))$. Conditions (12) and (13) insure that the time delays do not increase faster than time itself.

Mixed Nonlinear Lumped and Hyperbolic Distributed Parameter Systems

Setting $K = H = 0$, $\beta = 1$, and $r_1 = 1$, we obtain the mixed lumped and hyperbolic distributed system

$$\dot{x}(t) = f(x(t), z(1, t), t) + \xi(t) \quad (17)$$

$$z_t(r, t) = -M(r, t)z_r(r, t) + g(z(r, t), r, t) + \zeta(r, t) \quad (18)$$

$$y(t) = h(x(t), z(r_1^*, t), \dots, z(r_\gamma^*, t), t) + \eta(t) \quad (19)$$

subject to (4)–(6). Thus, (17)–(19) represents processes in which transportation lags are accompanied by phenomena such as dissipation of mass and energy, fluid mixing, and chemical reactions. In such cases, differential-difference

equations are inadequate in describing the system. The importance of this class of systems has been previously discussed by Hiratsuka and Ichikawa [24] and Aggarwal [25].

Nonlinear Systems with Functional Time Delays

The system (1)–(6) can be reduced to the following nonlinear lumped parameter system containing functional time delay:

$$\dot{x}(t) = f(x(t), x(t - \alpha_1), \dots, x(t - \alpha_\beta), t) + \int_0^{\alpha_{\max}} K_0(x(t - \alpha), \alpha, t) d\alpha + \xi(t) \quad (20)$$

$$y(t) = h(x(t), x(t - \alpha_1^*), \dots, x(t - \alpha_\gamma^*), t) + \int_0^{\alpha_{\max}} H_0(x(t - \alpha), \alpha, t) d\alpha + \eta(t) \quad (21)$$

$$x(t) = \phi(t), \quad -\alpha_{\max} \leq t \leq 0 \\ \alpha_{\max} = \max(\alpha_\beta, \alpha_\gamma^*) \quad (22)$$

where $0 < \alpha_1 < \dots < \alpha_\beta$ and $0 < \alpha_1^* < \dots < \alpha_\gamma^*$. This can be done by setting $g = \zeta = 0$, $b(x(t)) = x(t)$, (and hence, $n_2 = n_1$), $M(r, t) = \alpha_{\max}^{-1}$, $r_i = \alpha_i / \alpha_{\max}$, $r_j^* = \alpha_j^* / \alpha_{\max}$, $r = \alpha / \alpha_{\max}$, $K(z(r, t), r, t) = \alpha_{\max} K_0(z(r, t), \alpha_{\max} r, t)$, $H(z(r, t), r, t) = \alpha_{\max} H_0(z(r, t), \alpha_{\max} r, t)$, and $z(r, 0) = \phi(-r\alpha_{\max})$. Then $z(r, t) = x(t - \alpha_i)$, $z(r_j^*, t) = x(t - \alpha_j^*)$, and $z(r, t) = x(t - \alpha)$.

III. DERIVATION OF THE FILTER

The derivation of the filter for the system of (1)–(6) consists of two parts. First, we formulate the problem of fixed time smoothing and present the necessary condition for optimality in the form of a two-point boundary value problem. Second, we convert the smoothing problem into the filtering problem using a formulation based on differential sensitivities [26]. The technique has several desirable features: 1) no *a priori* assumptions regarding the form of the filter are required; 2) the exact interpretation of the so-called covariance matrices results; and 3) an indication of the form of the exact filter results. Although the derivation of the approximate nonlinear filter could have been carried out by invariant imbedding (for example, see Hwang *et al.* [12]), the present technique is less cumbersome. In this section we shall present the detailed derivation for the case of $K = H = 0$ in (1)–(6). We do this only for the convenience of the reader so as to avoid details which are more tedious than need be given. Because the form of the filter for the case of nonzero kernels, K and H , can be obtained so readily from that for $K = H = 0$, we merely present the filter in Section IV.

Statement of the Problem

Consider the system (1)–(6) with $K = H = 0$. The state estimation problem is: Given any fixed $T > 0$ and observations $y(t)$, $0 \leq t \leq T$, it is desired to estimate $x(t)$ and $z(r, t)$ for $0 \leq t \leq T$, $0 \leq r \leq 1$. This is the smoothing problem. The estimation criterion shall be to minimize

$$\Psi = \int_0^T \langle \dot{x} - f, R_0(t)(\dot{x} - f) \rangle dt \\ + \int_0^T \langle y - h, Q(t)(y - h) \rangle dt \\ + \int_0^T \left\{ \int_0^1 \int_0^1 \langle z_i(r, t) + M(r, t)z_r(r, t) \right. \\ \left. - g(z(r, t), r, t), R_1(r, s, t)(z_r(s, t) + M(s, t)z_s(s, t) \right. \\ \left. - g(z(s, t), s, t)) \rangle dr ds \right\} dt \quad (23)$$

where the matrices $R_0(t)$ and $Q(t)$ are symmetric positive-definite.

$R_1(r, s, t)$ is defined by

$$\int_0^1 R_1^+(r, \rho, t) R_1(\rho, s, t) d\rho = I\delta(r - s) \quad (24)$$

where $R_1^+(r, s, t)$ is a positive-definite, symmetric matrix: $R_1^+(r, s, t) = (R_1^+(s, r, t))^T$. $\delta(\cdot)$ is the Dirac delta function and I is the identity matrix. Although the weighting matrices $R_0(t)$ and $Q(t)$ and matrix $R_1^+(r, s, t)$ are only restricted to be symmetric positive-definite, and satisfy (24), they can be chosen to reflect the statistical properties of the stochastic variables $\xi(t)$, $\zeta(r, t)$, and $\eta(t)$ if statistical information about these errors is known.

We first reformulate this problem as an optimal control problem, i.e., it is desired to minimize

$$\psi_1 = \int_0^T \langle u(t), R_0(t)u(t) \rangle dt \\ + \int_0^T \langle y - h, Q(t)(y - h) \rangle dt \\ + \int_0^T \left\{ \int_0^1 \int_0^1 \langle v(r, t), R_1(r, s, t)v(s, t) \rangle dr ds \right\} dt \quad (25)$$

subject to the constraints

$$\dot{x}(t) = f(x(t), z(r_1, t), \dots, z(r_\beta, t), t) + u(t) \quad (26)$$

$$z_i(r, t) = -M(r, t)z_r(r, t) + g(z(r, t), r, t) + v(r, t) \quad (27)$$

$$z(0, t) = b(x(t)). \quad (28)$$

The necessary conditions for optimality corresponding to (25)–(28) are readily derived through adjoining (26)–(28) to the objective (25) by Lagrange multipliers $\lambda(t)$ and $\sigma(r, t)$ and then taking first variations. Only the results are presented here, where we use the carat $\hat{\cdot}$ to indicate the optimal values, and the notation (\cdot/T) in the arguments to denote the dependence of the optimal solution on the observation interval $[0, T]$. The optimal values of $\hat{x}(t/T)$ and $\hat{z}(r, t/T)$ result from the solution of the following two-point boundary value problem [27]:

$$\hat{x}_i(t/T) = \hat{f} - \frac{1}{2}R_0^{-1}(t)\hat{\lambda}(t/T) \quad (29)$$

$$\hat{z}_i(r, t/T) = -M\hat{z}_r + \hat{g} - \frac{1}{2} \int_0^1 R_1^-(r, s, t)\hat{\sigma}(s, t/T) ds \quad (30)$$

$$\hat{\lambda}_i(t/T) = 2\hat{h}_r^T Q(t)(y - \hat{h}) - \hat{f}_r^T \hat{\lambda}(t/T) \\ - \hat{b}_r^T M^T(0, t)\hat{\sigma}(0, t/T) \quad (31)$$

$$\begin{aligned} \hat{\sigma}_t(r,t/T) &= 2 \sum_{i=1}^{\gamma} \hat{h}_{z(r_i^*,t/T)}^T Q(t) (y - \hat{h}) \delta(r - r_i^*) \\ &\quad - \sum_{i=1}^{\beta} \hat{f}_{z(r_i,t/T)}^T \hat{\lambda}(t/T) \delta(r - r_i) \\ &\quad - \hat{g}_z^T \hat{\sigma}(r,t/T) - (M^T(r,t) \hat{\sigma}(r,t/T))_r \end{aligned} \quad (32)$$

$$\hat{\lambda}(0/T) = \hat{\lambda}(T/T) = 0 \quad (33)$$

$$\hat{\sigma}(r,0/T) = \hat{\sigma}(r,T/T) = 0 \quad (34)$$

$$\hat{z}(0,t/T) = b(\hat{x}(t/T)) \quad (35)$$

$$\hat{\sigma}(1,t/T) = 0 \quad (36)$$

where \hat{f} denotes $f(\hat{x}(t/T), \hat{z}(r_1,t/T), \dots, \hat{z}(r_\beta,t/T), t)$, etc. Equations (29)–(36) represent the boundary value problem which must be solved to produce the optimal least square smoothed estimates of $x(t)$ and $z(r,t)$ when data are given over $0 \leq t \leq T$. The optimal smoothing results for each of the special cases discussed in Section II can be determined from the appropriate simplification of these equations.

Differential Sensitivities

In the above two-point boundary value problem we can express the solutions $\hat{x}(t/T)$ and $\hat{z}(r,t/T)$ in terms of the Lagrange multipliers by

$$\hat{x}(t/T) = x[\hat{\lambda}(t/T), \hat{\sigma}(s,t/T)] \quad (37)$$

$$\hat{z}(r,t/T) = z[r, \hat{\lambda}(t/T), \hat{\sigma}(s,t/T)], \quad s \in [0,1]. \quad (38)$$

Let $\delta/\delta\hat{\sigma}$ denote the functional derivative and define the first-order differential sensitivity matrices P^{xx}, P^{xz}, P^{zx} , and P^{zz} by

$$P^{xx}(t/T) = -2 \frac{\partial \hat{x}(t/T)}{\partial \hat{\lambda}(t/T)} \quad (39)$$

$$P^{xz}(s,t/T) = -2 \frac{\partial \hat{x}(t/T)}{\partial \hat{\sigma}(s,t/T)} \quad (40)$$

$$P^{zx}(r,t/T) = -2 \frac{\partial \hat{z}(r,t/T)}{\partial \hat{\lambda}(t/T)} \quad (41)$$

$$P^{zz}(r,s,t/T) = -2 \frac{\partial \hat{z}(r,t/T)}{\partial \hat{\sigma}(s,t/T)}. \quad (42)$$

Then, using the chain rule of calculus, the partial derivatives of \hat{x} , \hat{z} , and \hat{z}_r with respect to T can be expressed as

$$\begin{aligned} \hat{x}_T(t/T) &= -\frac{1}{2} \left\{ P^{xx}(t/T) \hat{\lambda}_T(t/T) \right. \\ &\quad \left. + \int_0^1 P^{xz}(s,t/T) \hat{\sigma}_T(s,t/T) ds \right\} \end{aligned} \quad (43)$$

$$\begin{aligned} \hat{z}_T(r,t/T) &= -\frac{1}{2} \left\{ \int_0^1 P^{zx}(r,s,t/T) \hat{\sigma}_T(s,t/T) ds \right. \\ &\quad \left. + P^{zz}(r,t/T) \hat{\lambda}_T(t/T) \right\} \end{aligned} \quad (44)$$

$$\begin{aligned} \hat{z}_{rT}(r,t/T) &= -\frac{1}{2} \left\{ \int_0^1 P_r^{zz}(r,s,t/T) \hat{\sigma}_T(s,t/T) ds \right. \\ &\quad \left. + P_r^{zz}(r,t/T) \hat{\lambda}_T(t/T) \right\}. \end{aligned} \quad (45)$$

These equations describe the time evolution of the optimal solutions, \hat{x} and \hat{z} , as the length of the observation interval T varies.

Now let $q(t/T)$ be whatever we desire to estimate in the system, based on observations $y(\tau)$, $\tau \in [0,T]$, and denote the optimal estimate of $q(t/T)$ by $\hat{q}(t/T)$. Since we are interested in the optimal filter estimate, we seek $\hat{q}(T/T)$, and, in particular, the total derivative $d\hat{q}(T/T)/dT$. We note that

$$\frac{d\hat{q}(T/T)}{dT} = \hat{q}_t(t/T)|_{t=T} + \hat{q}_T(t/T)|_{t=T}$$

which we write for convenience as

$$\frac{d\hat{q}(T/T)}{dT} = \hat{q}_t(T/T) + \hat{q}_T(T/T). \quad (46)$$

Thus, the total derivative of the quantity $\hat{q}(T/T)$ is a sum of two terms, one representing the dynamics of the system, $\hat{q}_t(t/T)|_{t=T}$, and the second the updating of the estimate in the face of new observations, $\hat{q}_T(t/T)|_{t=T}$. This result was demonstrated for lumped parameter systems by Padmanabhan [26].

When \hat{q} is also a function of one or more spatial variables, $\hat{q}(r,s,t/T)$, then (46) becomes

$$\frac{\partial \hat{q}(r,s,T/T)}{\partial T} = \hat{q}_t(r,s,t/T)|_{t=T} + \hat{q}_T(r,s,t/T)|_{t=T} \quad (47)$$

which we write for convenience as

$$\frac{\partial \hat{q}(r,s,T/T)}{\partial T} = \hat{q}_t(r,s,T/T) + \hat{q}_T(r,s,T/T). \quad (48)$$

We emphasize that each term in (47), and hence (48), represents a *different* partial derivative. In particular, the left-hand side (LHS) of (47) and (48) is the analog to the total derivative in (46), whereas the right-hand side (RHS) of (47) and (48) consists of partial derivatives with respect to each of the arguments t and T in $(\cdot, t/T)$, respectively.

State Filter Equations

We now wish to derive the dynamical equations for $d\hat{x}(T/T)/dT$ and $\partial \hat{z}(r,T/T)/\partial T$ which represent the rate of change of the *filtered* estimates with T . Using (46) and (48), these can be expressed as

$$\frac{d\hat{x}(T/T)}{dT} = \hat{x}_t(T/T) + \hat{x}_T(T/T) \quad (49)$$

$$\frac{\partial \hat{z}(r,T/T)}{\partial T} = \hat{z}_t(r,T/T) + \hat{z}_T(r,T/T). \quad (50)$$

Equations (33) and (34) imply that

$$\frac{d\hat{\lambda}(T/T)}{dT} = \frac{\partial \hat{\sigma}(r, T/T)}{\partial T} = 0. \quad (51)$$

Using (46) and (48), (51) can be written

$$\hat{\lambda}_T(T/T) + \hat{\lambda}_t(T/T) = 0 \quad (52)$$

$$\hat{\sigma}_T(r, T/T) + \hat{\sigma}_t(r, T/T) = 0. \quad (53)$$

Then (31)–(34) and (52)–(53) give

$$\hat{\lambda}_T(T/T) = -2\hat{h}_x^T Q(T)(y - \hat{h}) \quad (54)$$

$$\hat{\sigma}_T(s, T/T) = -2 \sum_{i=1}^{\gamma} \hat{h}_x^T z_{(r_i^*, T/T)} Q(T)(y - \hat{h}) \delta(s - r_i^*). \quad (55)$$

Substituting (54) and (55) into (43) and (44), we obtain

$$\begin{aligned} \hat{x}_T(T/T) &= P^{xx}(T/T) \hat{h}_x^T Q(T)(y - \hat{h}) \\ &+ \sum_{i=1}^{\gamma} P^{xx}(r_i^*, T/T) \hat{h}_x^T z_{(r_i^*, T/T)} Q(T)(y - \hat{h}) \end{aligned} \quad (56)$$

$$\begin{aligned} \hat{z}_T(r, T/T) &= \sum_{i=1}^{\gamma} P^{zz}(r, r_i^*, T/T) \hat{h}_x^T z_{(r_i^*, T/T)} Q(T)(y - \hat{h}) \\ &+ P^{zz}(r, T/T) \hat{h}_x^T Q(T)(y - \hat{h}). \end{aligned} \quad (57)$$

On the other hand, (29), (30), (33), and (34) give

$$\hat{x}_t(T/T) = f \quad (58)$$

$$\hat{z}_t(r, T/T) = -M\hat{z}_r + \hat{g}. \quad (59)$$

Hence, (49), (50), and (56)–(59) constitute the state filter equations. The boundary condition for $\hat{z}(0, T/T)$ is

$$\hat{z}(0, T/T) = b(\hat{x}(T/T)). \quad (60)$$

Covariance Equations

We now need to derive the dynamic equations for the differential sensitivities, $P^{xx}(T/T)$, $P^{zz}(s, T/T)$, $P^{zx}(r, T/T)$, and $P^{zz}(r, s, T/T)$, to complete the specification of the filter. These equations are usually referred to as the covariance equations by analogy to the linear case, although they are not the true covariances in the nonlinear case. In order to derive these equations, we need the total derivatives with respect to T of the four differential sensitivities as in (49) and (50) for the state filter equations. As we know, each $\partial P(\cdot, T/T)/\partial T$ will be a sum of two terms, $P_t(\cdot, T/T)$ and $P_T(\cdot, T/T)$. For the general nonlinear case we are considering, it can be shown that $P_T(\cdot, T/T)$ involves the second-order differential sensitivities, and, likewise, the second-order differential sensitivities involve the *third-order* differential sensitivities etc. Thus, in general, it is not possible to close the system of equations. For this reason, we will approximate $\partial P(\cdot, T/T)/\partial T$ by $P_t(\cdot, T/T)$, enabling us to obtain a closed set of equations.

The basic approach is that we shall derive two expressions for each of the quantities,

$$\frac{\partial}{\partial t} [\hat{x}_T(t/T)] \quad \frac{\partial}{\partial t} [\hat{z}_T(r, t/T)] \quad (61)$$

and equate the two expressions for each of the quantities while setting $t = T$. Since each of the quantities above is a continuous function of t and T , we can write

$$\frac{\partial}{\partial t} [\hat{x}_T(t/T)] = \frac{\partial}{\partial T} [\hat{x}_t(t/T)] \quad (62)$$

$$\frac{\partial}{\partial t} [\hat{z}_T(r, t/T)] = \frac{\partial}{\partial T} [\hat{z}_t(r, t/T)]. \quad (63)$$

Substituting (29) in the RHS of (62) gives

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{x}_T(t/T)] &= f_x \hat{x}_T(t/T) + \int_0^1 \left\{ \sum_{i=1}^{\beta} f_{z(r_i, t/T)} \delta(r - r_i) \right\} \\ &\cdot \hat{z}_T(r, t/T) dr - \frac{1}{2} R_0^{-1}(t) \hat{\lambda}_T(t/T) \end{aligned} \quad (64)$$

which, with the help of (43) and (44), can be written as

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{x}_T(t/T)] &= -\frac{1}{2} \left[f_x P^{xx}(t/T) \right. \\ &+ \sum_{i=1}^{\beta} f_{z(r_i, t/T)} P^{xx}(r_i, t/T) + R_0^{-1}(r) \left. \right] \\ &\cdot \hat{\lambda}_T(t/T) - \frac{1}{2} \int_0^1 \left\{ \sum_{i=1}^{\beta} f_{z(r_i, t/T)} P^{zz}(r_i, s, t/T) \right. \\ &\left. + f_x P^{zz}(s, t/T) \right\} \hat{\sigma}_T(s, t/T) ds. \end{aligned} \quad (65)$$

On the other hand, using (43) we can write

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{x}_T(t/T)] &= -\frac{1}{2} \left[P_t^{xx}(t/T) \hat{\lambda}_T(t/T) \right. \\ &+ \int_0^1 P_t^{xz}(s, t/T) \hat{\sigma}_T(s, t/T) ds \left. \right] \\ &- \frac{1}{2} \left[P^{xx}(t/T) \frac{\partial}{\partial t} [\hat{\lambda}_T(t/T)] \right. \\ &\left. + \int_0^1 P^{zz}(s, t/T) \frac{\partial}{\partial t} [\hat{\sigma}_T(s, t/T)] ds \right] \end{aligned} \quad (66)$$

which gives us two expressions for the first quantity in (61).

To obtain two relations for the second quantity in (61) we first substitute (30) in the RHS of (63), giving

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{z}_T(r, t/T)] &= -M\hat{z}_{rT} + \hat{g}_{z(r, t/T)} \hat{z}_T(r, t/T) \\ &- \frac{1}{2} \int_0^1 R_1^+(r, s, t) \hat{\sigma}_T(s, t/T) ds \end{aligned} \quad (67)$$

which, with the help of (44) and (45), can be written as

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{z}_T(r, t/T)] &= -\frac{1}{2} \left[-M(r, t) P_r^{zz}(r, t/T) \right. \\ &+ g_z P^{zz}(r, t/T) \hat{\lambda}_T(t/T) \\ &- \frac{1}{2} \left[\int_0^1 \left\{ -M(r, t) P_r^{zz}(r, s, t/T) \right. \right. \\ &\left. \left. + \hat{g}_{z(r, t/T)} P^{zz}(r, s, t/T) \right. \right. \\ &\left. \left. + R_1^+(r, s, t) \right\} \hat{\sigma}_T(s, t/T) ds \right]. \end{aligned} \quad (68)$$

On the other hand, using (44) we can write

$$\begin{aligned} \frac{\partial}{\partial t} [\hat{z}_T(r,t/T)] = & -\frac{1}{2} \left[\int_0^1 P_i^{zz}(r,s,t/T) \hat{\sigma}_T(s,t/T) ds \right. \\ & \left. + P_i^{zz}(r,t/T) \hat{\lambda}_T(t/T) \right] \\ & - \frac{1}{2} \left[\int_0^1 P^{zz}(r,s,t/T) \frac{\partial}{\partial t} [\hat{\sigma}_T(s,t/T)] ds \right. \\ & \left. + P^{zz}(r,t/T) \frac{\partial}{\partial t} [\hat{\lambda}_T(t/T)] \right]. \end{aligned} \tag{69}$$

Now we equate (65) and (66) setting $t = T$. For the equality to hold the coefficients of $\hat{\lambda}_T(t/T)$, $\hat{\sigma}_T(s,t/T)$, and $\hat{\sigma}_T(0,t/T)$ must be zero at $t = T$. Doing so, we obtain [where it should be noted that some lengthy algebraic manipulations are required to evaluate the last two terms of (66)]

$$\begin{aligned} P_i^{zz}(T/T) = & \hat{f}_x P^{zz}(T/T) + P^{zz}(T/T) \hat{f}_x^T \\ & + \sum_{i=1}^{\beta} \hat{f}_{z(r_i, T/T)} P^{zz}(r_i, T/T) \\ & + \sum_{i=1}^{\beta} P^{zz}(r_i, T/T) \hat{f}_{z(r_i, T/T)}^T \\ & + P^{zz}(T/T) V^{zz}(T/T) P^{zz}(T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(T/T) V^{zi}(T/T) P^{zz}(r_i^*, T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r_i^*, T/T) V^{iz}(T/T) P^{zz}(T/T) \\ & + \sum_{i=1}^{\gamma} \sum_{j=1}^{\gamma} P^{zz}(r_i^*, T/T) V^{ij}(T/T) P^{zz} \\ & \cdot (r_j^*, T/T) + R_0^{-1}(T) \end{aligned} \tag{70}$$

$$\begin{aligned} P_i^{zz}(s, T/T) = & \hat{f}_x P^{zz}(s, T/T) + P^{zz}(s, T/T) \hat{g}_{z(s, T/T)}^T \\ & - P_s^{zz}(s, T/T) M^T(s, T) \\ & + \sum_{i=1}^{\beta} \hat{f}_{z(r_i, T/T)} P^{zz}(r_i, s, T/T) \\ & + P^{zz}(T/T) V^{zz}(T/T) P^{zz}(s, T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(T/T) V^{zi}(T/T) P^{zz}(r_i^*, s, T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r_i^*, T/T) V^{iz}(T/T) P^{zz}(s, T/T) \\ & + \sum_{i=1}^{\gamma} \sum_{j=1}^{\gamma} P^{zz}(r_i^*, T/T) V^{ij}(T/T) P^{zz} \\ & \cdot (r_j^*, s, T/T) \end{aligned} \tag{71}$$

$$P^{zz}(0, T/T) = P^{zz}(T/T) \hat{b}_x^T. \tag{72}$$

Similarly, equating (68) and (69) we obtain (again noting that the evaluation of the last two terms of (69) requires some algebraic labor)

$$\begin{aligned} P_i^{zz}(r, s, T/T) = & \hat{g}_{z(r, T/T)} P^{zz}(r, s, T/T) \\ & + P^{zz}(r, s, T/T) \hat{g}_{z(s, T/T)}^T \\ & - M(r, T) P_r^{zz}(r, s, T/T) \\ & - P_s^{zz}(r, s, T/T) M^T(s, T) \\ & + P^{zz}(r, T/T) V^{zz}(T/T) P^{zz}(s, T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r, T/T) V^{zi}(T/T) P^{zz}(r_i^*, s, T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r, r_i^*, T/T) V^{iz}(T/T) P^{zz}(s, T/T) \\ & + \sum_{i=1}^{\gamma} \sum_{j=1}^{\gamma} P^{zz}(r, r_i^*, T/T) V^{ij} \\ & \cdot (T/T) P^{zz}(r_j^*, s, T/T) \\ & + R_1^+(r, s, T) \end{aligned} \tag{73}$$

$$\begin{aligned} P_i^{zz}(r, T/T) = & P^{zz}(r, T/T) \hat{f}_x^T + \hat{g}_{z(r, T/T)} P^{zz}(r, T/T) \\ & - M(r, T) P_r^{zz}(r, T/T) \\ & + \sum_{i=1}^{\beta} P^{zz}(r, r_i, T/T) \hat{f}_{z(r_i, T/T)}^T \\ & + P^{zz}(r, T/T) V^{zz}(T/T) P^{zz}(T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r, r_i^*, T/T) V^{iz}(T/T) P^{zz}(T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r, T/T) V^{zi}(T/T) P^{zz}(r_i^*, T/T) \\ & + \sum_{i=1}^{\gamma} \sum_{j=1}^{\gamma} P^{zz}(r, r_i^*, T/T) V^{ij}(T/T) \\ & \cdot P^{zz}(r_j^*, T/T) \end{aligned} \tag{74}$$

$$P^{zz}(r, 0, T/T) = P^{zz}(r, T/T) \hat{b}_x^T. \tag{75}$$

The quantities V^{zz} , V^{iz} , V^{zi} , and V^{ij} are defined by

$$V^{zz}(T/T) = [\hat{h}_z^T Q(T)(y(T) - \hat{h})]_z \tag{76}$$

$$\begin{aligned} V^{iz}(T/T) = & [\hat{h}_{z(r_i^*, T/T)}^T Q(T)(y(T) - \hat{h})]_{z_i} \\ & i = 1, 2, \dots, \gamma \end{aligned} \tag{77}$$

$$\begin{aligned} V^{zi}(T/T) = & [\hat{h}_z^T Q(T)(y(T) - \hat{h})]_{z(r_i^*, T/T)}, \\ & i = 1, 2, \dots, \gamma \end{aligned} \tag{78}$$

$$\begin{aligned} V^{ij}(T/T) = & [\hat{h}_{z(r_i^*, T/T)}^T Q(T)(y(T) - \hat{h})]_{z(r_j^*, T/T)}, \\ & i, j = 1, 2, \dots, \gamma. \end{aligned} \tag{79}$$

The remaining boundary conditions for (70), (71), (73), and (74) can be obtained by differentiating (35) with respect to T ,

$$\begin{aligned} \frac{\partial \hat{z}(0, t/T)}{\partial \hat{\lambda}(t/T)} \hat{\lambda}_T(t/T) + \int_0^1 \frac{\delta \hat{z}(0, t/T)}{\delta \hat{\sigma}(s, t/T)} \hat{\sigma}_T(s, t/T) ds \\ = \hat{b}_x \frac{\partial \hat{x}(t/T)}{\partial \hat{\lambda}(t/T)} \hat{\lambda}_T(t/T) + \hat{b}_x \int_0^1 \frac{\delta \hat{x}(t/T)}{\delta \hat{\sigma}(s, t/T)} \hat{\sigma}_T(s, t/T) ds \end{aligned} \tag{80}$$

and then equating the coefficients of $\hat{\lambda}_T(t/T)$ and $\hat{\sigma}_T(s, t/T)$ to zero at $t = T$. The result is

$$P^{zz}(0, T/T) = \hat{b}_z P^{zz}(T/T) \quad (81)$$

$$P^{zz}(0, s, T/T) = \hat{b}_z P^{zz}(s, T/T). \quad (82)$$

The entire filter is summarized in Table I. In the column of initial conditions, $\hat{x}(0/0)$ and $\hat{z}(r, 0/0)$ represent our best initial guesses of x_0 and $z_0(r)$. The initial conditions $P^{zz}(0/0)$, $P^{zz}(s, 0/0)$, $P^{zz}(r, 0/0)$ and $P^{zz}(r, s, 0/0)$ are basically arbitrary. In the linear, white noise case it can be shown that

$$P^{zz}(T/T) = E[(x(T) - \hat{x}(T/T))(x(T) - \hat{x}(T/T))^T] \quad (83)$$

$$P^{zz}(r, T/T) = E[(x(T) - \hat{x}(T/T))(z(r, T) - \hat{z}(r, T/T))^T] \quad (84)$$

$$P^{zz}(r, T/T) = E[(z(r, T) - \hat{z}(r, T/T))(x(T) - \hat{x}(T/T))^T] \quad (85)$$

$$P^{zz}(r, s, T/T) = E[(z(r, T) - \hat{z}(r, T/T))(z(s, T) - \hat{z}(s, T/T))^T]. \quad (86)$$

These relations may be used as a guide in choosing $P^{zz}(0/0), \dots, P^{zz}(r, s, 0/0)$.

Discussion of the Filter

The exact equations for the four covariance matrices are of the form

$$\frac{dP(T/T)}{dT} = P_r(T/T) + P_T(T/T) \quad (87)$$

where P can denote P^{xx} , P^{xz} , P^{zx} , or P^{zz} .

We noted earlier that we would neglect the second terms on the RHS of these equations. Let us give some indication as to how these neglected terms might be calculated. Employing the chain rule, we have for $P_{T^{zz}}(r, t/T)$, for example,

$$P_{T^{zz}}(r, t/T) = \frac{\partial P^{zz}(r, t/T)}{\partial \hat{\lambda}(t/T)} \hat{\lambda}_T(t/T) + \int_0^1 \frac{\partial P^{zz}(r, t/T)}{\partial \hat{\sigma}(v, t/T)} \hat{\sigma}_T(v, t/T) dv. \quad (88)$$

The terms

$$\frac{\partial P^{zz}(r, t/T)}{\partial \hat{\lambda}(t/T)} \quad \frac{\partial P^{zz}(r, t/T)}{\partial \hat{\sigma}(v, t/T)}$$

are $n_1 \times n_2 \times n_1$ and $n_1 \times n_2 \times n_2$ matrices, respectively, which are the second-order differential sensitivities. Thus, the neglected terms in (87) involve second-order differential sensitivities, which, in turn, depend on third-order differential sensitivities, etc. As with other nonlinear stochastic problems in mathematics, the exact solution of the nonlinear filtering problem is unavailable due to a closure problem. In the linear, white noise case it can be shown that the second and higher order differential sensitivities are identically zero. It is the neglect of the second term on the RHS of the equations of the type (87) that

TABLE I
FILTER FOR THE SYSTEM OF (1)-(6) WITH $K = H = 0$

	Equations	Initial Conditions (arbitrary)	Boundary Conditions
Estimates			
$\hat{x}(T/T)$	(49), (56), (58)	$\hat{x}(0/0)$	None
$\hat{z}(r, T/T)$	(50), (57), (59)	$\hat{z}(r, 0/0)$	(60)
First Order Differential Sensitivities			
$P^{zz}(T/T)$	(70)	$P^{zz}(0/0)$	None
$P^{zz}(s, T/T)$	(71)	$P^{zz}(s, 0/0)$	(72)
$P^{zz}(r, T/T)$	(74)	$P^{zz}(r, 0/0)$	(81)
$P^{zz}(r, s, T/T)$	(73)	$P^{zz}(r, s, 0/0)$	(75), (82)

constitutes the approximation inherent in the nonlinear filter. Inclusion of higher order differential sensitivities increases the complexity of the filter, the implementation of which would be considerably more difficult computationally. As with all approximate nonlinear filters, their validity can only be assessed through computer simulation.

IV. FILTERING IN NONLINEAR SYSTEMS DESCRIBED BY FUNCTIONAL DIFFERENTIAL EQUATIONS

The method of derivation of Section III can be used in the case when K and H are nonzero in (1) and (3). By replacing h by $h + \int_0^1 H dr$ and f by $f + \int_0^1 K dr$ in Ψ in (23), the filter equations for system model (1)-(3) can readily be obtained. The filter in this case is summarized below. The state filter is

$$\begin{aligned} \frac{d\hat{x}(T/T)}{dT} = & \hat{f} + \int_0^1 K[\hat{z}(\theta, T/T), \theta, T] d\theta \\ & + P^{zz}(T/T) \hat{h}_x^T Q(T) \hat{\Phi}(T/T) \\ & + \sum_{i=1}^{\gamma} P^{zz}(r_i^*, T/T) \hat{h}_{z(r_i^*, T/T)}^T Q(T) \hat{\Phi}(T/T) \\ & + \int_0^1 P^{zz}(\theta, T/T) \hat{H}_{z(\theta, T/T)}^T(\theta) Q(T) \hat{\Phi}(T/T) d\theta \end{aligned} \quad (89)$$

$$\begin{aligned} \frac{d\hat{z}(r, T/T)}{dT} = & -M\hat{z}_r + \hat{g} + \sum_{i=1}^{\gamma} P^{zz}(r, r_i^*, T/T) \hat{h}_{z(r_i^*, T/T)}^T \\ & \cdot Q(T) \hat{\Phi}(T/T) \\ & + P^{zz}(r, T/T) \hat{h}_x^T Q(T) \hat{\Phi}(T/T) \\ & + \int_0^1 P^{zz}(r, \theta, T/T) \hat{H}_{z(\theta, T/T)}^T \\ & \cdot (\theta) Q(T) \hat{\Phi}(T/T) d\theta \end{aligned} \quad (90)$$

$$\hat{z}(0, T/T) = b(\hat{x}(T/T)) \quad (91)$$

where

$$\begin{aligned} \hat{\Phi}(T/T) = & y(T) - h(\hat{x}(T/T), \hat{z}(r_1^*, T/T), \dots, \\ & \cdot \hat{z}(r_\gamma^*, T/T), T) - \int_0^1 H(\hat{z}(\theta, T/T), \theta, T) d\theta. \end{aligned} \quad (92)$$

The covariance equations are (corresponding to (70), (71), (73), and (74) for $K = H = 0$)

TABLE II
FILTER FOR THE SYSTEM (1)–(6) WITH $K \neq 0, H \neq 0$

Equations	Initial Conditions (arbitrary)	Boundary Conditions
Estimates		
$\hat{x}(T/T)$	$\hat{x}(0/0)$	None
$\hat{z}(r,T/T)$	$\hat{z}(r,0/0)$	(91)
First Order Differential Sensitivities		
$P^{zz}(T/T)$	$P^{zz}(0/0)$	None
$P^{zz}(s,T/T)$	$P^{zz}(s,0/0)$	(72)
$P^{zz}(r,T/T)$	$P^{zz}(r,0/0)$	(81)
$P^{zz}(r,s,T/T)$	$P^{zz}(r,s,0/0)$	(75), (82)

$$P_i^{zz}(T/T) = [\text{RHS of (70)}] + \int_0^1 \hat{K}_{z(\theta,T/T)}(\theta) P^{zz}(\theta,T/T) d\theta + \int_0^1 P^{zz}(\theta,T/T) \hat{K}_{z(\theta,T/T)}^T(\theta) d\theta + W^{zz}(T/T) \quad (93)$$

$$P_i^{zz}(s,T/T) = [\text{RHS of (71)}] + \int_0^1 \hat{K}_{z(\theta,T/T)}(\theta) P^{zz}(\theta,s,T/T) d\theta + W^{zz}(s,T/T) \quad (94)$$

$$P_i^{zz}(r,T/T) = [\text{RHS of (74)}] + \int_0^1 P^{zz}(r,\theta,T/T) \hat{K}_{z(\theta,T/T)}^T(\theta) d\theta + W^{zz}(r,T/T) \quad (95)$$

$$P_i^{zz}(r,s,T/T) = [\text{RHS of (73)}] + W^{zz}(r,s,T/T) \quad (96)$$

where the term $V(T/T)$ will be defined below (the definition differs slightly from that in Section III). The terms $W^{\mu\lambda}(\cdot, T/T)$ are defined by

$$W^{\mu\lambda}(\cdot, \cdot, T/T) = \int_0^1 P^{\mu x}(\cdot, T/T) V^{zz}(\xi, T/T) P^{z\lambda}(\xi, \cdot, T/T) d\xi + \int_0^1 P^{\mu z}(\cdot, \xi, T/T) V^{zz}(\xi, T/T) P^{z\lambda}(\cdot, T/T) d\xi + \sum_{i=1}^{\gamma} \int_0^1 P^{\mu z}(\cdot, r_i^*, T/T) V^{iz}(\xi, T/T) P^{z\lambda}(\xi, \cdot, T/T) d\xi + \sum_{i=1}^{\gamma} \int_0^1 P^{\mu z}(\cdot, \xi, T/T) V^{zi}(\xi, T/T) P^{z\lambda}(r_i^*, \cdot, T/T) d\xi + \int_0^1 \int_0^1 P^{\mu z}(\cdot, \xi, T/T) V_1^{zz}(\xi, \nu, T/T) P^{z\lambda}(\nu, \cdot, T/T) d\xi d\nu + \int_0^1 \int_0^1 P^{\mu z}(\cdot, \xi, T/T) V_2^{zz}(\xi, \nu, T/T) P^{z\lambda}(\xi, \cdot, T/T) d\xi d\nu + \int_0^1 P^{\mu z}(\cdot, \xi, T/T) V_3^{zz}(\xi, T/T) P^{z\lambda}(\xi, \cdot, T/T) d\xi \quad (97)$$

where $\mu = x$ or z and $\lambda = x$ or z . The unspecified left argument of each term is r if $\mu = z$ and does not exist if $\mu = x$. The unspecified right argument of each term is s if $\lambda = z$ and does not exist if $\lambda = x$.

In this case the V matrices are defined as follows (the dependence on T is suppressed for convenience):

$$V^{zz} = [\hat{h}_x^T Q \hat{\Phi}]_x \quad V^{zi} = [\hat{h}_x^T Q \hat{\Phi}]_{z(r_i^*)}$$

$$V^{zz}(\theta) = -\hat{h}_x^T Q \hat{H}_{z(\theta)}(\theta) \quad V^{iz} = [\hat{h}_{z(r_i^*)}^T Q \hat{\Phi}]_x$$

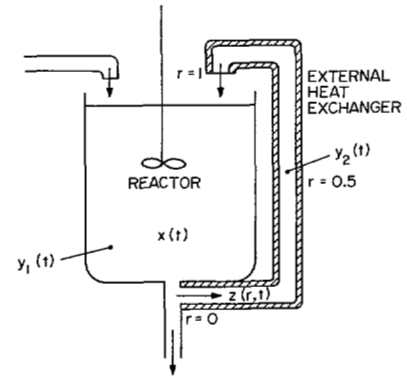


Fig. 2. Well-stirred chemical reactor with external heat exchanger.

$$V^{ij} = [\hat{h}_{z(r_i^*)}^T Q \hat{\Phi}]_{z(r_i^*)} \quad V^{iz}(\theta) = -\hat{h}_{z(r_i^*)}^T Q \hat{H}_{z(\theta)}(\theta)$$

$$V^{zz}(\theta) = -\hat{H}_{z(\theta)}^T(\theta) Q \hat{h}_x \quad V^{zi}(\theta) = -\hat{H}_{z(\theta)}^T(\theta) Q \hat{h}_{z(r_i^*)}$$

$$V_1^{zz}(\theta, \nu) = -\hat{H}_{z(\theta)}^T(\theta) Q \hat{H}_{z(\nu)}(\nu)$$

$$V_2^{zz}(\theta, \nu) = -[\hat{H}_{z(\theta)}^T(\theta) Q \hat{H}(\nu)]_{z(\theta)}$$

$$V_3^{zz}(\theta) = [\hat{H}_{z(\theta)}^T(\theta) Q (y - \hat{h})]_{z(\theta)} \quad (98)$$

The filter in this case is summarized in Table II.

V. ESTIMATION OF TEMPERATURE IN A RECYCLE REACTOR

An important problem in the process industries is the estimation of the temperature distribution in a chemical reactor. A rather common configuration for the execution of highly exothermic reactions is a system consisting of a well-stirred chemical reactor a portion of the output from which is recycled through a heat exchanger back to the reactor. By such a design, the temperature of the well-stirred reactor, in which the reaction is primarily carried out, is controlled by recycling a fixed fraction of the effluent through the heat exchanger. Such a system is shown in Fig. 2.

Let us consider the problem of estimating the temperatures, $x(t)$ and $z(r,t)$, in the well-stirred reactor and heat exchanger, respectively, based on corrupted measurement of the reactor temperature $x(t)$ and the temperature at the midpoint of the heat exchanger, $z(0.5,t)$. We consider the case in which the chemical reaction is zero order, that is the rate of reaction is independent of the concentration of reactant. Under such conditions, we need not include differential equations for species concentration, only the two temperatures, $x(t)$ and $z(r,t)$. If $x(t)$ and $z(r,t)$ are made dimensionless, the dynamic behavior of the system is governed by

$$\dot{x}(t) = -0.05x + 0.2z(1,t) - 0.001 \left[\exp\left(\frac{20x}{1+x}\right) - 1 \right]$$

$$z_t(r,t) = -z_r(r,t) + 0.1 - 0.001 \exp\left(\frac{20z}{1+z}\right),$$

$$0 \leq r \leq 1$$

$$z(0,t) = x(t), \quad t \geq 0$$

where the parameters have been chosen to be typical of

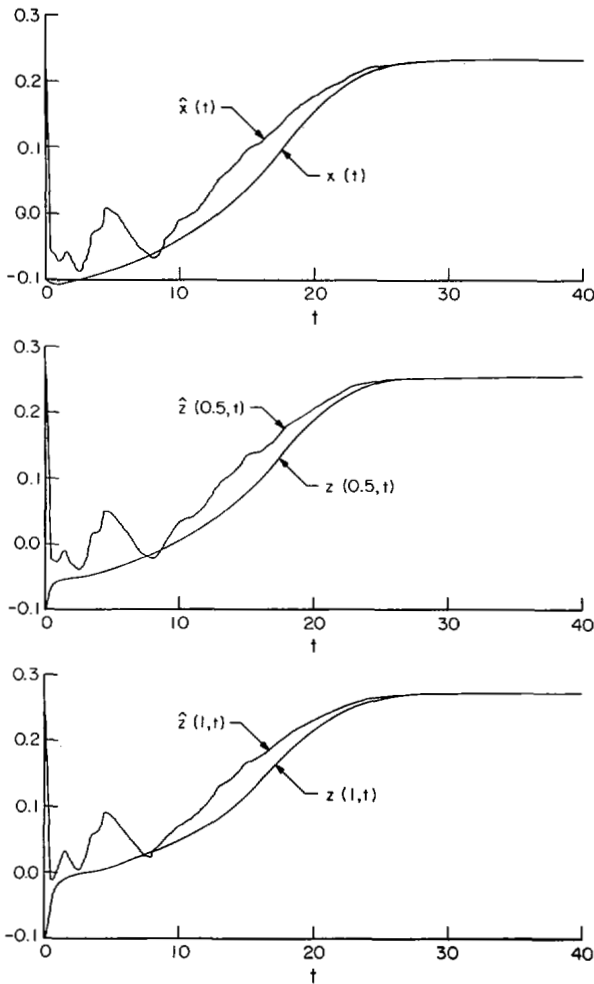


Fig. 3. Comparison of actual and estimated temperatures in reactor, at midpoint and exit of heat exchanger, $x(t)$, $z(0.5,t)$, $z(1,t)$, respectively.

those that might be encountered in practice. The non-linearity of the system arises from the exponential dependence of the rate of a chemical reaction on temperature. Note that the chemical reaction proceeds not only in the well-stirred reactor but also in the heat exchanger, as indicated by the exponential terms in both state equations. The observations are given by

$$y_1(t) = x(t) + \eta_1(t) \quad y_2(t) = z(0.5,t) + \eta_2(t)$$

on the basis of which we desire to estimate $x(t)$ and $z(r,t)$. In chemical reactor problems of this type it is often acceptable to neglect dynamical noise [28]. The observation errors η_1 and η_2 arise because of turbulent eddies passing the temperature probes.

For the purpose of numerical simulation we assume that the true, but unknown, initial states are $x(0) = z(r,0) = -0.1$. The observation errors were generated by $\eta_i(t) = 0.3G(0,1)$, $i = 1,2$, where $G(0,1)$ is a normally distributed random variable with zero mean and unit standard deviation produced by a random number generator. We assume that the level and character of this noise is unknown. Numerical experiments were conducted with

the following parameter values: $\hat{x}(0) = \hat{z}(r,0) = 0.25$, $P^{zz}(0) = 2.0$, $P^{zz}(s,0) = 2.0$, $P^{zz}(r,s,0) = 2.0$, and $Q = 1.0$. Note that because we do not have any *a priori* information on the initial states or the noise, these parameters are arbitrary. Numerical implementation of the state and filter equations was carried out using a standard finite difference scheme and required about 40 seconds on an IBM 370/155 for an experimental time of $t = 40$.

The filter performance is shown in Fig. 3, where the true states, $x(t)$, $z(0.5,t)$, $z(1,t)$, and their estimates $\hat{x}(t)$, $\hat{z}(0.5,t)$, and $\hat{z}(1,t)$ are compared. Experiments with different initial guesses and values of Q confirmed the robustness of the filter for the relatively significant measurement noise considered here. These results point to advantages in the use of a filter in a feedback control system for chemical reactors of the type considered here.

VI. CONCLUSIONS

In this paper we have derived approximate smoothing and filtering equations for a general class of nonlinear functional differential equations, (1)–(6). From these results we have obtained filters for the following classes of systems:

- 1) nonlinear lumped parameter systems with multiple constant and time-varying delays;
- 2) nonlinear systems of coupled ordinary and hyperbolic partial differential equations; and
- 3) nonlinear systems having functional time delays.

With the exception of the linear filters of Kwakernaak [18] and Koivo [20], which are encompassed within our general results, the filters developed herein are new.

REFERENCES

- [1] A. V. Balakrishnan and J. L. Lions, "State estimation for infinite dimensional systems," *J. Comput. Syst. Sci.*, vol. 1, pp. 391–403, 1967.
- [2] P. L. Falb, "Infinite-dimensional filtering: The Kalman–Bucy filter in Hilbert space," *Inform. Contr.*, vol. 11, pp. 102–137, 1967.
- [3] F. E. Thau, "On optimal filtering for a class of linear distributed parameter systems," *Trans. ASME (J. Basic Eng.)*, series D91, pp. 173–178, 1969.
- [4] S. G. Tzafestas and J. M. Nightingale, "Optimal filtering, smoothing, and prediction in linear distributed-parameter systems," *Proc. Inst. Elec. Eng.*, vol. 115, pp. 1207–1212, 1968.
- [5] —, "Concerning optimal filtering theory of linear distributed-parameter systems," *Proc. Inst. Elec. Eng.*, vol. 115, pp. 1737–1742, 1968.
- [6] —, "Boundary and volume filtering of linear distributed-parameter systems," *Electron. Lett.*, vol. 5, pp. 199–200, 1969.
- [7] —, "Maximum-likelihood approach to the optimal filtering of distributed parameter systems," *Proc. Inst. Elec. Eng.*, vol. 116, pp. 1085–1093, 1969.
- [8] H. J. Kushner, "Filtering for linear distributed parameter systems," *SIAM J. Contr.*, vol. 8, pp. 346–359, 1970.
- [9] J. S. Meditch, "Least-square filtering and smoothing for linear distributed-parameter systems," *Automatica*, vol. 7, pp. 315–322, 1971.
- [10] J. S. Meditch, "On state estimation for distributed parameter systems," *J. Franklin Inst.*, vol. 290, no. 1, p. 49, 1970.
- [11] J. H. Seinfeld, G. R. Gavalas, and M. Hwang, "Nonlinear filtering in distributed parameter systems," *J. Dynamic Syst., Meas. Contr.*, vol. 93G, pp. 157–163, 1971.
- [12] M. Hwang, J. H. Seinfeld, and G. R. Gavalas, "Optimal least-square filtering and interpolation in distributed parameter systems," *J. Math. Anal. Appl.*, vol. 39, no. 1, pp. 49–74, 1972.
- [13] S. G. Tzafestas, "On optimum distributed-parameter filtering

and fixed-interval smoothing for colored noise," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 448-458, Aug. 1972.

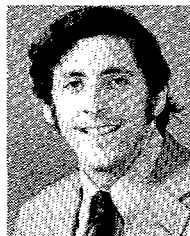
- [14] Y. Sakawa, "Optimal filtering in linear distributed parameter systems," *Int. J. Contr.*, vol. 16, no. 1, pp. 115-127, 1972.
- [15] S. G. Tzafestas, "Bayesian approach to distributed parameter filtering and smoothing," *Int. J. Contr.*, vol. 15, no. 2, pp. 273-295, 1972.
- [16] A. Bensoussan, *Filtrage Optimal des Systemes Lineaires*. Paris: Dunod, 1971.
- [17] G. Lamont and K. S. P. Kumar, "State estimation in distributed parameter systems via least squares and invariant embedding," *J. Math. Anal. Appl.*, vol. 38, pp. 588-606, 1972.
- [18] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 169-173, Apr. 1967.
- [19] L. Padmanabhan, "On filtering in delay systems with continuous dynamics and discrete-time observations," presented at the 1972 Joint Automatic Control Conf., Paper 29-1, pp. 826-831.
- [20] A. J. Koivo, "Optimal estimation for linear stochastic systems described by functional differential equations," *Inform. Contr.*, vol. 19, pp. 232-245, 1971.
- [21] A. Lindquist, "A theorem on duality between estimation and control for linear stochastic systems with time delay," *J. Math. Anal. Appl.*, vol. 37, pp. 516-536, 1972.
- [22] V. Shukla and M. D. Srinath, "Optimal filtering in linear distributed parameter systems with multiple time delays," *Int. J. Contr.*, vol. 16, pp. 673-687, 1972.
- [23] R. F. Curtain and P. L. Falb, "Stochastic differential equations in Hilbert space," *J. Differential Equations*, vol. 10, no. 3, pp. 412-430, 1971.
- [24] S. Hiratsuka and A. Ichikawa, "Optimal control of systems with transportation lags," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 237-247, June 1969.
- [25] J. K. Aggarwal, "Feedback control of linear systems with distributed delay," in *Proc. IFAC Symp. Dist. Parameter Systems*, Banff, June 1971.
- [26] L. Padmanabhan, "A new approach to recursive estimation in continuous dynamical systems," *Chem. Eng. J.*, vol. 1, pp. 232-240, 1970.
- [27] T. K. Yu, "Optimal filtering for systems governed by coupled ordinary and partial differential equations," Ph.D. dissertation, California Inst. Technol., Pasadena, 1973.
- [28] J. H. Seinfeld, G. R. Gavalas, and M. Hwang, "Control of nonlinear stochastic systems," *Ind. Eng. Chem. Fundamentals*, vol. 8, pp. 257-262, 1969.



Thomas K. Yu was born in Chungking, China, in 1945. He received the B.S. degree in electrical engineering from the University of California, Los Angeles, in 1966 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1967 and 1973, respectively.

His summer working experiences were with the U.S. Naval Undersea Center, Pasadena, Calif., T.R.W., Redondo Beach, Calif., and

the Jet Propulsion Laboratory, Pasadena, Calif. He is now with the Aerospace Group, Missile Systems Division of the Hughes Aircraft Co. in Canoga Park, Calif.



John H. Seinfeld was born in Elmira, New York on August 3, 1942. He received the B.S. degree in chemical engineering from the University of Rochester, Rochester, N.Y., in 1964, and the Ph.D. degree in chemical engineering from Princeton University, Princeton, N.J., in 1967.

Since 1967 he has been at the California Institute of Technology, Pasadena, where he is now Associate Professor of Chemical Engineering. His interests are in estimation

and filtering theory and in the mathematical modeling and control of urban air pollution. He is co-author of *Numerical Solution of Ordinary Differential Equations* (New York: Academic, 1971) and *Mathematical Methods in Chemical Engineering*, vol. III—*Process Modeling, Estimation, and Identification* (Englewood Cliffs, N.J.: Prentice-Hall, 1974). He is Associate Editor of *Chemical Engineering Communications* and serves on the Editorial Board of *Environmental Science and Technology*.

Dr. Seinfeld received the Donald P. Eckman Award of the American Automatic Control Council in 1970, and was a recipient of a Camille and Henry Dreyfus Foundation Teacher-Scholar Grant in 1972.



W. Harmon Ray was born in Washington, D.C. on April 4, 1940. He received the B.A. and B.S. Ch.E. degrees from Rice University, Houston, Tex., in 1962 and 1963, respectively, and the Ph.D. degree in chemical engineering from the University of Minnesota, Minneapolis, in 1966.

From 1966 to 1970 he taught chemical engineering at the University of Waterloo, Waterloo, Ont., Canada. Since 1970 he has been with the State University of New York

at Buffalo, where he is now Professor of Chemical Engineering. During the 1973-1974 academic year, Professor Ray was on sabbatical leave as a Guggenheim Fellow in Belgium and Germany, where he is collaborating on on-line estimation and control projects. His research interests are in process modeling, optimization, and control, and is co-author of *Process Optimization* (New York: Wiley, 1973).

Dr. Ray is the recipient of the 1969 Donald P. Eckman Award of the American Automatic Control Council.